# קובץ – buildModel

פונק' buildTheModel

הפונק' נקראת כאשר לוחצים על כפתור הbuild. הפונקציה מייצרת אובייקט מטיפוס dataPreperation ואובייקט נוסף מטיפוס calassifierBuilder. לאחר מכן, הפונק' קוראת לפונק' אשר מפורטות בהמשך המסמך.

פונק' buildStructure

הפונק' קוראת מהדיסק את קובץ הנstructure, בודקת תקינות של הקובץ, ומוציאה עבור כל attribute האם הוטא קטגוריאלי או נומרי במקרה של attribute קטגוריאלי, היא מחלצת את הקטגוריות השונות האפשריות עבור הattribute.

פונק' readTrain

הפונק' קוראת את קובץ ה-train אל תוך dataset.

# מחלקת DataPreperaion

בכדי ליצור מופע מטיפוס DataPreperation, יש צורך בdataset של ה-Train. הattributes, והקטגוריות עבור כל attribute קטגוריאלי.

פונק' handleNans

עבור attribute קטגוריאלי, משבצת את הערך שמופיע הכי הרבה. עבור attribute נומרי היא משבצת את הממוצע.

פונק' discretization

מחלקת לbins את ה attribute-ים הנומריים. בנוסף, הפונק' שומרת את הbins שהוגדרו, בכדי לאפשר חלוקה לbins לקובץ הtest.

# מחלקת classifierBuilder

בכדי ליצור מופע מטיפוס DataPreperation, יש צורך בdataset של ה-Train. הattributes, והקטגוריות עבור כל attribute קטגוריאלי, bins, ועבור כל attribute את הbins הספציפיים לו.

הבנאי של המחלקה קורא לפונק' (המפורטות מטה) שמייצרות עבורו את המידע הנדרש עבור ביצוע חיזוי.

פונק' calcNforeachClass

עבור כל סיווג אפשרי, הפונקציה בודקת מה ההסתברות להיות בסיווג הזה, בהינתן הdataset של הtrain.

פונק' classify

מקבלת את הpath לtest, וקוראת לפונק' המפורטות מטה.

פונק' getTest

הפונק' קוראת את הtest, מתקן Nanים לפי ההגדרות, ומחלק לbins עפ"י הbins ששמרנו והתקבלו בבנאי המחלקה.

פונק' calc

הפנוק' עוברת על הtest ומפעילה עבור כל שורה את הפונקציה של חישוב הnaive bayes ולפונקציה אשר מדפיסה לקובץ הoutput.

פונק' naiveBayesForRow

זאת הפונק אשר בפועל מבצעת את החישוב ואת הסיווג של כל שורה לאחד הסיווגים האפשריים.

פונק printRes

הפונק' מקבלת את הpath של התקייה, את מספר השורה בתוך הtest, ואת הסיווג.

הפונק' מדפיסה את התוצאה לתוך קובץ הoutput.